

Eyewitnesses memory for faces in actual criminal cases: An archival analysis of positive facial composites

Renan Benigno Saraiva. University of Portsmouth
Goiara Mendonça de Castilho. Universidade de Brasília
Raiane Nunes Nogueira. Universidade de Brasília
Letícia de Amorim Mota Coelho. Universidade de Brasília
Luciana Carvalho Pulschen Alarcão. Universidade de Brasília
Jade Lage. Universidade de Brasília

Abstract

Facial composites are crucial in the criminal justice system. In this archival study, we investigated the assumption that the success of facial composites depends partly on variables related to the crime, which either impairs or facilitates mnemonic processes. When a facial composite is successful in taking an offender to court it is sometimes archived as a positive facial composite, including a photo of the culprit and information about the crime. A total of 88 positive facial composites were investigated. The accuracy of facial composites was tested as a function of five variables related to the crime: type of crime, presence of weapon, retention interval, exposure duration, and disguise. Participants judged the resemblance of the perpetrators' photo with their correspondent facial composite. The results pointed out that only exposure duration was significantly associated with facial composites accuracy. Possible implications and future directions for research using archived facial composites are discussed.

Keywords: eyewitness testimony; facial composites; cognitive psychology; memory; face perception.

Resumo

Memória de testemunhas oculares para faces em casos criminais: Uma análise de arquivo de retratos falados positivos. Retratos falados são de extrema importância no sistema de justiça criminal. Neste estudo, nós investigamos a hipótese de que o sucesso de retratos falados depende parcialmente de variáveis relacionadas ao crime, que podem tanto facilitar quanto dificultar processos mnemônicos na elaboração de retratos falados. Quando um retrato falado obtém sucesso em levar um infrator à corte ele é comumente arquivado como um retrato falado positivo, incluindo uma foto do culpado e informações sobre o crime. Um total de 88 retratos falados positivos foram investigados. A acurácia de cada retrato falado foi testada de acordo com cinco variáveis relacionadas ao crime: tipo de crime, presença de arma, intervalo de retenção, tempo de exposição e disfarce. Participantes avaliaram a semelhança entre a foto do culpado e seu respectivo retrato falado. Os resultados apontaram que somente o tempo de exposição foi associado com a acurácia dos retratos falados. Possíveis aplicações e futuras pesquisas usando retratos falados reais são discutidos.

Palavras-chave: testemunho; retratos falados; psicologia cognitiva; memória; percepção de faces.

Resumen

Memoria de testigos oculares para rostros en casos criminales: Un análisis de archivo de retratos hablados positivos. Retratos hablados son de suma importancia en el sistema de justicia criminal. En este estudio, se investigó la hipótesis de que el éxito de los retratos hablados depende en parte de las variables relacionadas con el delito, que pueden facilitar o dificultar los procesos mnemotécnicos en el desarrollo de los retratos hablados. Cuando un retrato hablado tiene éxito en llevar un transgresor a la justicia este es archivado como un retrato hablado positivo, incluyendo una foto del culpable y la información sobre el crimen. Se estudió un total de 88 retratos hablados positivos. La precisión de cada imagen se puso a prueba en función de cinco variables relacionadas con el crimen: tipo de delito, la presencia de armas, intervalo de retención, tiempo de exposición y el encubrimiento. Participantes evaluaron la similitud entre la imagen del culpable y su respectivo retrato hablado. Los resultados mostraron que sólo el tiempo de exposición se asoció con la precisión de retratos hablados. Se discuten las posibles aplicaciones y futuras investigaciones utilizando retratos hablados reales.

Palabras clave: testigos oculares; retratos hablados; psicología cognitiva; memoria; percepción de rostros.

Statements provided by eyewitnesses are critical in criminal investigations lacking significant physical evidences. Initially those reports can include general descriptions of the suspect appearance, providing information about sex, race and approximate height and age. In some cases, an investigator works with the eyewitness to construct a facial composite of the offender, resulting in an image intended to resemble the facial characteristics of the perpetrator (McQuiston-Surrett, Topp, & Malpass, 2006; Zahradnikova, Duchovicova, & Schreiber, 2016). Those images are then sent to police officers working in the field, in order to narrow the investigation of suspects and assist on the identification of perpetrators. In such cases, it necessary and desirable that the facial composites are constructed as accurate as possible, maximizing the chances of bringing a perpetrator to justice. In some situations, the person convicted for a particular crime has many physical similarities with a previously constructed facial composite, named from now on as positive facial composite. Unfortunately, many findings show that eyewitness memory can be inaccurate, and many different variables have been found to impair eyewitness performance (Houston, Hope, Memon, & Don Read, 2013; Lindsay, Ross, Read, & Toglia, 2013; Wells & Hasel, 2007; Wells & Olson, 2003). In this archival study, we aim to investigate positive facial composites constructed by eyewitnesses of real crimes, analyzing how facial composites quality is influenced by variables related to the crime.

Facial composites are created with the purpose of reducing the number of suspects and to help identify a perpetrator, relying on someone's ability to recognize his or her face (Zahradnikova et al., 2016). Accurate composites work well as an investigative tool, but unfortunately an extensive literature shows that eyewitness memory can also be unreliable (Lindsay, Mansour, Kalmet, Bertrand, & Melsom, 2011; Lindsay et al., 2013; Memon, Mastroberardino, & Fraser, 2008), and that composites generated by eyewitnesses are usually not easily recognized (Davies & Valentine, 2013; Wells & Hasel, 2007). A great number of individual and situational factors account for eyewitness identification accuracy, often categorized as *estimator* or *system* variables (Wells, 1978; Wells & Olson, 2003). Estimator variables refer to factors that are beyond the control of the criminal system, such as duration of the crime, characteristics of the culprit and visibility conditions during the event. System variables are factors that can

somehow be controlled by the criminal system, such as instructions given by agents during the construction of facial composites. In this study, we highlight five estimator variables that are reported along with the positive facial composites investigated, specifically: type of crime, presence of weapon, retention interval, exposure duration, and use of disguise.

Estimator Variables

There seems to be a lack of studies that systematically compare different types of crimes, but many studies have investigated the role of arousal and stress elicited by criminal events in eyewitness memory (Block, Greenberg, & Goodman, 2009; Deffenbacher, 1983; Deffenbacher, Bornstein, Penrod, & McGorty, 2004). The exact role of emotion and stress in eyewitness performance is still an unresolved issue. Some authors emphasize the attentional orientation function of the arousing stimuli (Deffenbacher et al., 2004), and others emphasize the defensive function of emotional reactions (Christianson, 1992; Christianson, 2014; Kim, Park, & Lee, 2014). Nevertheless, most of the literature points out that witnesses have an improved performance for central aspects of emotional negative events, and an inferior performance for peripheral details of the event (Edelstein, Alexander, Goodman, & Newton, 2004; Reisberg & Heuer, 2004, 2007). Although definitions of central and peripheral details are usually confounded in the literature, in most studies central information is described as elements directly associated with the source of the emotional stress (e.g. characteristics of the offender), while peripheral information is externally and not associated with the source of stress (e. g. background scenery).

The presence or absence of a weapon is also relevant when witnessing a crime, affecting witnesses' attention and consequently identification accuracy. The weapon focus effect shows that when a weapon is present during the crime, the performance of eyewitnesses' memory is usually impaired (Fawcett, Russell, Peace, & Christie, 2013; Kramer, Buckhout, & Eugenio, 1990; Pickel, Ross, & Truelove, 2006). The presence of a weapon not only increases arousal in the moment of the crime, but also creates a strong attentional distraction, which in turn increases memory accuracy for details related to the weapon, while decreasing accuracy on other details, such as physical characteristics of the perpetrator (Carlson, Dias, Weatherford, & Carlson, 2017; Fawcett et al., 2013; Flowe, Hope, & Hillstrom, 2013; Hope & Wright, 2007; Steblay, 1992). However,

some archival studies did not support the weapon focus effect, either showing a complete absence of effect (Behrman & Davey, 2001), or only marginally significant results (Tollestrup, Turtle, & Yuille, 1994).

The time-lapse between codification and retrieval of information has long been associated with loss of details in memory (Ebbinghaus, 1885/1964), and the same pattern emerges in eyewitness literature. Sauer, Brewer, Zweck and Weber (2010), for example, found that identification responses made after three to seven weeks after encoding were less likely to be correct than responses provided immediately after witnessing the event. Additionally, this same study shows that retention interval affects the confidence-accuracy relationship, so that longer retention interval increases overconfidence. Part of the detrimental effect of retention interval on memory has been associated with source monitoring, so that longer retention intervals increase the inclusion of new and misleading information to the initial memory (Johnson, Hashtroudi, & Lindsay, 1993). Furthermore, a meta-analysis of 53 facial memory studies revealed a strong association between longer retention intervals and poorer recall performance (Deffenbacher, Bornstein, McGorty, & Penrod, 2008).

Exposure duration has also been shown to affect memory performance, so that the time eyewitnesses have to observe a perpetrator positively correlates with accuracy in subsequent memory tasks (Bornstein, Deffenbacher, Penrod, & McGorty, 2012; Reynolds & Pezdek, 1992; Shapiro & Penrod, 1986). In a study conducted by Memon, Hope and Bull (2003), for example, it was found that eyewitness had an improved identification performance in situations with longer exposure duration (45s) when compared with a short exposure (12s). Additionally, Palmer, Brewer, Weber and Nagesh (2013), found a small but significant effect of exposure duration on eyewitness accuracy, with more correct responses in a 90s exposure condition (63.8%), and less correct responses in a 5s exposure condition (50.7%). A meta-analysis on the effects of exposure duration in facial identification accuracy also supports that correct identifications are significantly higher in longer exposure durations (Bornstein et al., 2012).

Several studies indicate a rather counter-intuitive finding concerning exposure duration, which states that false alarm rates are also more prevalent as exposure duration increases. Read (1995) argues that this may occur mainly because witnesses consider meta-memorial information about the length of exposure when

deciding whether a face was seen before. Memory for details that have been seen for a longer period would be more easily available, which increase the rates of false identifications and false claims. Those findings suggest that bigger exposure duration can have both facilitating (higher identification accuracy) and detrimental (more false identifications) effects on eyewitnesses' identification.

Research in eyewitness and face recognition literature greatly considers the impact of disguising individual facial features in memory accuracy. Some studies point out that masking eyes, hair or mouth significantly impairs latter recognition of the suspect, highlighting the role of specific features in memory consolidation of faces (Lindsay et al., 2011; Shapiro & Penrod, 1986; Terry, 1993). Brewer, Weber and Semmler (2005) argue that disguises can impair witnesses' memories in two non-exclusive ways, affecting the quantity of features codified and/or increasing the perceived difficulty of the identification task. Mansour et al. (2012) provide two more explanations, the first related to coding specificity, in which details of material are more easily remembered if the encoded information is presented in the same way as it was shown initially. The second explanation concerns attentional resources, so that disguises could generate either a cognitive load feeling or call the witness attention to the disguise, instead of other relevant facial features.

Methodological Issues

This study focuses on an unconventional method in the eyewitness literature, using facial composites from real police cases as study subject. It is not surprising that most studies in eyewitness literature are based on experimental procedures, in which participants are exposed to live crime simulations or videos, engaging later in some sort of identification task. However, the small number of case studies has been criticized as a lack of external validity in eyewitness research, emphasizing the necessity of a wide variety of methods to investigate eyewitness memory issues (Tollestrup et al., 1994; Yuille, 1993, Yuille, 2013). Real case studies allow the testing of a range of variables that would be too complex to study altogether in controlled experiments (Behrman & Davey, 2001; Horry, Halford, Brewer, Milne, & Bull, 2014; Horry, Memon, Wright, & Milne, 2012). But there are many empirical limitations and concerns on the use of archived facial composites.

A major concern in case studies is that there is no irrefutable assurance that the convicted person

is indeed the same person who committed the crime (Horry et al., 2014). An archived facial composite, for example, could be paired with another person erroneously, which would distort any analysis using this data. However, some measures can be taken to alleviate or completely avoid this problem, using other extrinsic incriminating evidence of the cases, such as confessions, fingerprints and DNA samples (Tollestrup et al., 1994). In order to minimize this issue, we selected files that had at least two other extrinsic incriminating evidences in addition to the facial composites. One could also be concerned that facial composites are used only in high profile cases, biasing analysis using real composites. That is not the case in the current study, as facial composites are used in a wide range of criminal cases in the jurisdiction where the study was conducted.

Another important limitation consists on the restrict quality range of facial composites, given that usually only composites of convicted persons are archived. The composites used in this study had to exceed a certain quality standard before they were used by the police, resulting in a rather restricted range of composite quality. This issue limits the interpretability of the effects found, so that the inferences made from real case files are limited to facial composites with moderate or good quality. However, despite of these methodological issues, archival studies have been pointed out as a meritorious complementary approach in the investigation of eyewitnesses' identification accuracy (Behrman & Davey, 2001; Fahsing, Ask, & Granhag, 2004; Horry et al., 2014). The use of real case scenarios or real police files is highly needed to improve the ecological validity of eyewitnesses' memory theoretical models (Yuille, 2013).

Predictions

According to the reviewed literature, we first hypothesized that facial composites of crimes with higher emotional arousal would be more accurate, given the increased attention to central details in stressful events. Based on the weapon focus effect, we predicted that facial composites would have poorer resemblance with the actual perpetrators in crimes using weapons. We also hypothesized that longer retention intervals would have a detrimental effect on the quality of facial composites, and that longer exposure would result in more reliable facial composites. Finally, we expected that facial composites of disguised criminals would have a poorer accuracy when compared with facial composites of non-disguised criminals.

Method

Descriptive Information of Facial Composites

We analyzed forensic files provided by a police department specialized in the construction of facial composites. The police department is located in a large urban area and creates from 15 to 30 composites per month. The facial composites are constructed using a feature selection system, and the whole procedure is conducted on Photoshop with a specialized officer that creates the composite with the eyewitness. An initial interview is conducted with the eyewitness to obtain overall information about the crime and the perpetrator physical characteristics. After the interview, the eyewitness select face features that best match the eyewitness memory of the perpetrator, and finally the officer modifies the image using Photoshop tools and following eyewitness instructions. An example of facial composite created with this method is showed in Figure 1.



Figure 1. Example of a Positive Facial Composite Used in the Study.

Eighty-eight files with positive facial composites were gathered, produced in investigations from 2009 to 2012. Each file consisted of: a) a facial composite created by a witness or victim, b) a photo of the convicted person, and c) details of the crime. The details of the crime included physical descriptions of the perpetrator, such as gender, race and approximate height and age, along with the type of crime, crime duration, presence of weapon, retention interval and whether the declarant was a victim or witness of the crime. The most documented crime was robbery ($n = 25$, 29%), and other cases included homicide ($n = 7$, 9%), residential

burglary ($n = 17$, 21%), rape ($n = 20$, 24%), and others ($n = 19$, 21%). Moreover, most of the crimes had a weapon involved (65%). Exposure duration ranged from five minutes to four hours, with an average of 58.8 minutes ($SD = 1.12$). In average, witnesses constructed the facial composite nine days after the crime ($SD = 25.3$), ranging from immediate testimonies to reports provided 154 days after the crime. Only two of the facial composites were from female perpetrators. Finally, 74 facial composites we constructed by victims (88.1%), and five facial composites were constructed by bystanders.

For each facial composite, there was a similarity rate given by eyewitnesses, provided on a scale from 1 to 10 according to the similarity of the facial composite in respect to the perpetrator's face. In the Criminal Justice System studied only facial composites with similarity rates higher than five are sent to police stations, thus all positive facial composites analyzed in this study had rates higher than five (with $M = 7.08$, $SD = 1.03$). Such rates seems useful as an assessment of witnesses' confidence, however it should not be considered an entirely reliable indicator of accuracy of the facial composite (Leippe & Eisenstadt, 2014). This rate is based mainly on a metacognitive effort, in which the witness compares the facial composite with the face he/she remembers without the chance of making direct comparisons. Therefore, we proceeded with a method to obtain a more trustful measure of the accuracy of the facial composites.

Evaluation of the Accuracy of Facial Composites

Participants. The sample consisted of 126 participants (50.4% female) aged between 17 and 21 years ($M = 20.0$, $SD = 2.16$), who took part in an evaluation task of the facial composites. Participants did not receive any reward for participation. This study received a favorable ethical review (CAAE: 36968214.1.0000.5540).

Materials and procedure. The 88 pairs of positive facial composites and real perpetrator photo were separated into six different blocks of images - four with 15 facial composites and two with 14 facial composites. Each participant received only one of the six blocks of images, and were instructed to judge the similarity of the first image (facial composite) in respect to the second image (photo of the perpetrator). Therefore, 21 participants evaluated each facial composite.

Initially all pairs of images were presented to the participant, one after the other (for 3 seconds each) in a computer screen. There was a 2 seconds interval between the presentations of each pair of images. This was performed in order to provide the participants a bigger picture of all facial composites, preventing anchorage effects. After seeing all images, the evaluation step was initiated. At this moment, the participants received a questionnaire to rate the image pairs for overall appearance, and for specific characteristics (eyes, mouth, ears, face shape and hair). Each characteristic needed to be judged in a scale ranging from 1 - not similar at all - to 10 - completely similar. The first pair was shown again and remained displayed in the screen until the participant finished the first evaluation. After the participant finished the first evaluation, he/she could continue to the next pair, with unlimited time to evaluate all the pairs in the block.

An average of all participants' evaluation was used as a new similarity index for each one of the facial composites. Therefore, we had a more objective similarity rate for the overall face appearance and for specific facial features (eyes, mouth, ears, face shape and hair), for each facial composite. Data from the two facial composites of female perpetrators was removed from analysis, to maintain congruence of distinctive gender facial features. Other three facial composites were also excluded due to very poor image quality, resulting in a total of 83 facial composites for analysis.

Results

The main descriptive statistics of the positive facial composites similarity rates are summarized in Table 1. It is important to keep in mind that eyewitnesses rate was provided in a scale from 1 to 10 with undefined limits, while participants used a scale from 1 - not similar at all - to 10 - completely similar. Despite of this consideration, participants average rate ($M = 4.34$) was considerably smaller than the rate that eyewitnesses gave to the facial composites ($M = 7.08$); $t(81) = 15.72$, $p < 0.001$, $d = 2.51$), revealing a possible discrepancy between eyewitnesses confidence and its actual accuracy. The absence of significant correlation between the eyewitnesses and participants rates further confirmed this observation ($r = -0.01$, $p = 0.45$). As for the specific facial features, the hair had the highest similarity average ($M = 5.26$), followed by the nose ($M = 4.93$), and mouth ($M = 4.84$).

Table 1. Similarity Scores Given by Eyewitnesses of the Crimes (ES) and by Participants of the Study (PS).

	ES	PS Whole face	PS Eyes	PS Nose	PS Mouth	PS Hair	PS Face shape
Mean	7.08	4.34	4.50	4.93	4.84	5.26	4.28
Standard deviation	1.03	1.14	1.28	1.28	1.18	1.63	1.36
Minimum	5.0	1.85	1.70	2.05	1.36	2.10	1.40
Maximum	9.0	6.85	6.78	7.78	7.09	8.23	7.45

Note. Eyewitnesses provided rates in a scale that ranged from 1 to 10, without defined limits. Participants provided rates in a scale from 1 – nothing similar at all to 10 – completely similar.

Type of Crime and Weapon Focus

From ten different types of documented crimes, only three of them had sufficient number of observations for analysis, specifically: residential burglary ($n = 17$), robbery ($n = 25$), and rape ($n = 20$). Facial composites of robberies had the highest similarity rates ($M = 4.42$, $SD = 1.07$), followed by rape ($M = 4.40$, $SD = 1.21$), and residential burglary ($M = 4.10$, $SD = 1.13$). However no significant differences were found between those means in one-way ANOVA, $F(9,73) = 0.75$, $p = 0.66$, $\eta^2 p = 0.014$. In addition, no difference was found in the similarity rates for crimes with weapons ($M = 4.34$, $SD = 1.12$) or without weapons ($M = 4.41$, $SD = 1.18$), $t(81) = 0.35$, $p = 0.72$, $d = 0.06$. Victims of crimes created composites with lower similarity rates ($M = 4.31$, $SD = 1.07$) when compared to bystanders ($M = 4.97$, $SD = 2.18$), but this difference was also not significant in an one tailed independent t test, $t(75) = 1.13$, $p = 0.26$, $d = 0.38$.

Retention Interval

Outliers in the retention interval variable were transformed in order to reduce their influence in normality, assigning their values to the maximum limit of this variable. This procedure greatly reduced the kurtosis on this variable (from 23.8 to 1.17). We found no significant correlation between retention interval and the overall similarity rates of facial composites given by participants ($r = -0.02$, $p = 0.40$). Similarity rates for facial composites reproduced up to three days after the crime, for example, had a 4.37 ($SD = 1.04$) mean, while facial composites made after three days or more had 4.29 ($SD = 1.07$) similarity mean. The same pattern of results was found in similarity rates of specific facial features.

Exposure Duration

Most crimes analyzed ranged from five to thirty minutes (44%), and maximum of four hours. We found that overall similarity of facial composites was positively correlated with duration of the crime ($r = 0.22$, $p < 0.01$), with an increased effect when controlling by the type of crime ($r = 0.27$, $p = 0.03$). However, we also found that the actual eyewitnesses rate of similarity was negatively correlated with duration of the crime, so that eyewitnesses of longer crimes tended to give lower similarity rates ($r = -0.17$, $p = 0.03$).

Disguise

Exactly 34.5% of the facial composites analyzed had some kind of disguise, mostly hats (77%). The average of similarity rates from facial composites with disguises ($M = 4.60$, $SD = 0.94$) was higher than the ones without disguises ($M = 4.21$, $SD = 1.22$), but such difference was not statistically significant $t(81) = 1.48$, $p = 0.14$, $d = 0.36$.

Discussion

We investigated how the facial composites accuracy could be influenced by variables related to the crime, evaluating archived positive facial composites of real cases. Our main predictions were that: a) Facial composites of crimes with high emotional arousal would be more accurate when compared to low emotional arousal, given the increased attention to central details during stressful events; b) Facial composites would have a poorer resemblance with the actual perpetrators in crimes with weapons; c) Longer retention intervals would have a detrimental effect on the quality of facial composites; d) Longer exposure durations would result in more reliable facial composites; and e) Facial composites would be less accurate when the criminal had a disguise during the crime.

We found that the similarity rate of facial composites of three different crimes (residential burglary, robbery and rape) were almost the same. Although this result does not comprise many other types of crimes, it is initial evidence that quality of facial composites might not be substantially associated with the type of crime witnessed. Such result contradicts our hypothesis that eyewitnesses would perform differently depending on the type of crime, due to differences in emotional intensities (Attwood, Penton-Voak, Burton, & Munafò, 2013; Hancock, Burke, & Frowd, 2011; Reisberg & Heuer, 2004, 2007). However, there is still an extensive controversy

about effects of stress in eyewitnesses' memory, with many evidences for positive, negative or null effects (Christianson, 1992; Deffenbacher, 1983; Deffenbacher et al., 2004). It seems that many of these differences could be explained by the use of different study designs, stimulus intensities and consistent definitions of stress and arousal (Deffenbacher et al., 2004).

We did not find evidences supporting the weapon focus effect, given that similarity rates of facial composites for crimes with or without weapons were almost the same. It was expected that in crimes with weapons witnesses would produce more inaccurate facial composites, as was confirmed in previous studies with a variety of weapons (Carlson et al., 2017; Fawcett et al., 2013; Kramer et al., 1990; Pickel et al., 2006). Although this is a well-established effect in laboratory conditions, some archival studies do not seem to support it (Behrman & Davey, 2001; Tollestrup et al., 1994). Crime seriousness and exposure duration are the two most common explanations for this inconsistency, given that in laboratory conditions the presence of a weapon is merely staged (Cutler & Penrod, 1995).

Longer retention intervals did not have a detrimental effect on the overall quality of facial composites. Many evidences confirm that eyewitnesses' accuracy decreases as the interval between crime and identification test increases (Deffenbacher et al., 2008). However, it is still possible that such effect is diminished in real cases due to other factors. Witnesses of real crimes often repeat many times what they have seen, reviving the codified information and minimizing the detrimental effect of retention interval (Ebbesen & Rienick, 1998). Furthermore, evidences suggest that crime seriousness and emotional impact can moderate the effects of retention interval on eyewitnesses' memory, so that longer intervals have less effect in eyewitnesses' memory related to crimes with high emotional intensities (Deffenbacher et al., 2004).

We found that longer exposure duration had a beneficial effect on facial composites quality, the same pattern of results found in laboratory studies (Memon et al., 2003; Read, 1995; Reynolds & Pezdek, 1992). In our study, this effect stood out among the other variables tested, given that only exposure duration reached a statistically significant relationship with facial composites similarity rates given by participants. The relation between exposure duration and memory accuracy is well established in facial recognition literature, although it is still an unexplored issue in facial

composites reproduction (Palmer et al., 2013; Shapiro & Penrod, 1986).

It was assumed that disguises, such as sunglasses or accessories covering the hair, would have a detrimental effect on eyewitnesses' accuracy (Hockley, Hemsworth, & Consoli, 1999). However, our results did not support this prediction. Facial composites with disguises actually received higher similarity rates, but without a statistically significant difference. Facial composites with disguises may have received better similarity rates because the very disguises prevented the participants to make a complete comparison between all facial features on the images, consequently giving higher rates based only on the remaining features.

Many other important variables in eyewitnesses' research were not addressed in this study, mainly because the files used did not have the necessary additional information. Cross-racial effects could not be tested, for example, because the perpetrator race is provided, but the eyewitnesses' race is not (Meissner & Brigham, 2001). Moreover, suggestibility and conformity effects are relevant in eyewitness reports, but are difficult to access in real case scenarios (Gabbert, Memon, & Allan, 2003; Zaragoza & Lane, 1994). Different composite systems could also modify the effects of estimator variables in eyewitnesses' performance, especially on the comparison of feature and holistic based systems (Frowd et al., 2011). It is then important to consider that our results contain many different sources of noise due to variables that were not possible to control, a common issue in archival studies that is not easily overcome (Horry et al., 2014). Other limitations in this study include the fact that all composites originated from the same place, which imply that results may not generalize to different contexts. Also, we did not have access or control about specific aspects of the construction of each facial composite, although all of them were constructed using the same guidelines.

Most of the predictions made in this study based on the current literature were not confirmed. A first interpretation of this outcome might suggest that most findings in laboratorial eyewitness studies still fail to generalize to naturalistic contexts, an issue also present in previous real case and archival studies (Behrman & Davey, 2001; Horry et al., 2014; Tollestrup et al., 1994). However, it is important to consider that effects found in laboratory studies can apply to real eyewitness situations, but in archival studies they might eventually lose power due to complex interactions between variables

(Yuille, 2013). Another reasonable explanation to the many null effects found in this study is related to the limited variability of the quality of facial composites used. All composites used in this study are from cases in which someone was convicted, and they had to exceed a certain quality standard before being used in the Criminal Justice System. As a result, the range of composite quality is restricted, with an absence of low quality facial composites. However, overcoming this limitation is not easy, given that low quality composites are often discarded or cannot be easily paired with the real perpetrator.

The investigation of eyewitnesses' accuracy in facial composite reconstruction is still an emerging area. Many estimator variables were already identified as relevant in the accuracy of eyewitnesses in identification procedures such as line-ups and photospreads (Wells & Hasel, 2007; Wells & Olson, 2003). However, only few studies tried to test these findings in facial composites reconstruction. Most studies in this topic focused on different kinds of facial composites systems or in facial features (Davies & Valentine, 2013; Frowd et al., 2005; Frowd, Bruce, McIntyre, & Hancock, 2007), but the scarcity of studies testing crime-related variables remains. We highlight the increasing necessity for archival studies to test these effects, which combined with laboratory studies could provide a wider comprehension on eyewitnesses' memory for crimes.

References

- Attwood, A. S., Penton-Voak, I. S., Burton, A. M., & Munafò, M. R. (2013). Acute anxiety impairs accuracy in identifying photographed faces. *Psychological Science*, 24(8), 1591-1594. doi: 10.1177/0956797612474021
- Behrman, B. W., & Davey, S. L. (2001). Eyewitness identification in actual criminal cases: an archival analysis. *Law and Human Behavior*, 25(5), 475-491. doi: 10.1023/A:1012840831846
- Block, S. D., Greenberg, S. N., & Goodman, G. S. (2009). Remembrance of eyewitness testimony: effects of emotional content, self-relevance, and emotional tone. *Journal of Applied Social Psychology*, 39(12), 2859-2878. doi: 10.1111/j.1559-1816.2009.00553.x
- Bornstein, B. H., Deffenbacher, K. A., Penrod, S. D., & McGorty, E. K. (2012). Effects of exposure time and cognitive operations on facial identification accuracy: a meta-analysis of two variables associated with initial memory strength. *Psychology, Crime & Law*, 18(5), 473-490. doi: 10.1080/1068316X.2010.508458
- Brewer, N., Weber, N., & Semmler, C. (2005). Eyewitness identification. In N. Brewer & K. D. Williams (Eds.), *Psychology and law: an empirical perspective* (pp. 177-221). New York, NY: Guilford.
- Carlson, C. A., Dias, J. L., Weatherford, D. R., & Carlson, M. A. (2017). An investigation of the weapon focus effect and the confidence-accuracy relationship for eyewitness identification. *Journal of Applied Research in Memory and Cognition*, 6(1), 82-92. doi: 10.1016/j.jarmac.2016.04.001
- Christianson, S. A. (1992). Emotional stress and eyewitness memory: a critical review. *Psychological Bulletin*, 112(2), 284-309. doi: 10.1037/0033-2909.112.2.284
- Christianson, S. A. (2014). *The handbook of emotion and memory: research and theory*. Psychology Press.
- Cutler, B. L., & Penrod, S. D. (1995). *Mistaken identification: the eyewitness, psychology, and the law*. Cambridge, England: Cambridge University Press.
- Davies, G., & Valentine, T. (2013). Facial composites: forensic utility and psychological research. In R. C. L. Lindsay, D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Handbook of eyewitness psychology: memory for people* (pp. 59-83). Mahwah, NJ: Erlbaum.
- Deffenbacher, K. A. (1983). The influence of arousal on reliability of testimony. In S. M. A. Lloyd-Boystock & B. R. Clifford (Eds.), *Evaluating witness evidence* (pp. 235-251). Chichester, England: Wiley
- Deffenbacher, K. A., Bornstein, B. H., Penrod, S. D., & McGorty, E. K. (2004). A meta-analytic review of the effects of high stress on eyewitness memory. *Law and Human Behavior*, 28(6), 687-706. doi: 10.1007/s10979-004-0565-x
- Deffenbacher, K. A., Bornstein, B. H., McGorty, E. K., & Penrod, S. D. (2008). Forgetting the once-seen face: estimating the strength of an eyewitness's memory representation. *Journal of Experimental Psychology: Applied*, 14(2), 139-150. doi: 10.1037/1076-898X.14.2.139
- Ebbesen, E. B., & Rienick, C. B. (1998). Retention interval and eyewitness memory for events and personal identifying attributes. *Journal of Applied Psychology*, 83, 745-762. doi: 10.1037/0021-9010.83.5.745
- Ebbinghaus, H. (1964). *Memory: a contribution to experimental psychology*. New York, NY: Dove. (Original work published in 1885)
- Edelstein, R. B., Alexander, K. W., Goodman, G. S., & Newton, J. W. (2004). Emotion and eyewitness memory. In D. Reisberg & P. Hertel (Eds.), *Memory and emotion* (pp. 308-346). New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195158564.003.0010
- Fahsing, I. A., Ask, K., & Granhag, P. A. (2004). The man behind the mask: accuracy and predictors of eyewitness offender descriptions. *Journal of Applied Psychology*, 89(4), 722-729. doi: 10.1037/0021-9010.89.4.722
- Fawcett, J. M., Russell, E. J., Peace, K. A., & Christie, J. (2013). Of guns and geese: a meta-analytic review of the 'weapon focus' literature. *Psychology, Crime & Law*, 19(1), 35-66. doi: 10.1080/1068316X.2011.599325
- Flowe, H. D., Hope, L., & Hillstrom, A. P. (2013). Oculomotor examination of the weapon focus effect: does a gun automatically engage visual attention?. *Plos One*, 12(8), 1-7. doi: 10.1371/journal.pone.0081011
- Frowd, C. D., Bruce, V., McIntyre, A., & Hancock, P. (2007). The relative importance of external and internal features of facial composites. *British Journal of Psychology*, 98(1), 61-77. doi: 10.1348/000712606X104481
- Frowd, C. D., Carson, D., Ness, H., Richardson, J., Morrison, L., McInaghlan, S., & Hancock, P. (2005). A forensically valid comparison of facial composite systems. *Psychology, Crime & Law*, 11(1), 33-52. doi: 10.1080/10683160310001634313
- Frowd, C. D., Hancock, P. J. B., Bruce, V., Skelton, F. C., Atherton, C., Nelson, L., ... Sendrea, G. (2011). Catching more offenders with

- EvoFIT facial composites: lab research and police field trials. *Global Journal of Human Social Science*, 11, 46-58. Retrieved from <http://socialscienceresearch.org/index.php/GJHSS/article/view/165>
- Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: can eyewitnesses influence each other's memories for an event?. *Applied Cognitive Psychology*, 17(5), 533-543. doi: 10.1002/acp.885
- Hancock, P. J., Burke, K., & Frowd, C. D. (2011). Testing facial composite construction under witness stress. *International Journal of Bio-Science and Bio-Technology*, 3(3), 6571. Retrieved from http://www.sersc.org/journals/IJBSBT/vol3_no3/6.pdf
- Hockley, W. E., Hemsworth, D. H., & Consoli, A. (1999). Shades of the mirror effect: recognition of faces with and without sunglasses. *Memory and Cognition*, 27(1), 128-38. doi:10.3758/BF03201219
- Hope, L., & Wright, D. (2007). Beyond unusual? Examining the role of attention in the weapon focus effect. *Applied Cognitive Psychology*, 21(7), 951-961. doi: 10.1002/acp.1307
- Horry, R., Halford, P., Brewer, N., Milne, R., & Bull, R. (2014). Archival analyses of eyewitness identification test outcomes: what can they tell us about eyewitness memory? *Law and Human Behavior*, 38(1), 94-108. <https://doi.org/10.1037/lhb0000060>
- Horry, R., Memon, A., Wright, D. B., & Milne, R. (2012). Predictors of eyewitness identification decisions from video lineups in England: a field study. *Law and Human Behavior*, 36(4), 257-265. doi: 10.1037/h0093959
- Houston, K. A., Hope, L., Memon, A., & Don Read, J. (2013). Expert testimony on eyewitness evidence: in search of common sense. *Behavioral Sciences & the Law*, 31(5), 637-651. doi: 10.1002/bsl.2080
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3-28. doi: 10.1037/0033-2909.114.1.3
- Kim, K., Park, K. K., & Lee, J. H. (2014). The influence of arousal and expectation on eyewitness memory in a virtual environment. *Cyberpsychology, Behavior and Social Networking*, 17(11), 709-713. doi: 10.1089/cyber.2013.0638
- Kramer, T. H., Buckhout, R., & Eugenio, P. (1990). Weapon focus, arousal, and eyewitness memory: attention must be paid. *Law and Human Behavior*, 14(2), 167-184. doi: 10.1007/BF01062971.
- Leippe, M. R., & Eisenstadt, D. (2014). Eyewitness confidence and the confidence-accuracy relationship in memory for people. In R. C. L. Lindsay, D. F. Ross., J. D. Read, & M. P. Toglia (Eds.), *The Handbook of Eyewitness Psychology: Volume II: Memory for People* (Vol. 2, pp. 377-425). Psychology Press.
- Lindsay, R. C. L., Mansour, J. K., Kalmel, N., Bertrand, M. I., & Melsom, L. (2011). Face perception and recognition in eyewitness memory. In A. Calder, G. Rhodes, M. Johnson, J. Haxby, & J. Keane (Eds.), *The Oxford handbook of face perception* (pp. 307-328). Oxford, England: Oxford University Press.
- Lindsay, R. C. L., Ross, D. F., Read, J. D., & Toglia, M. P. (Eds.) (2013). *The handbook of eyewitness psychology: memory for people* (Vol. 2). Psychology Press.
- Mansour, J. K., Beaudry, J. L., Bertrand, M. I., Kalmel, N., Melsom, E. I., & Lindsay, R. C. (2012). Impact of disguise on identification decisions and confidence with simultaneous and sequential lineups. *Law and Human Behavior*, 36(6), 513-526. doi: 10.1037/h0093937
- McQuiston-Surrett, D., Topp, L. D., & Malpass, R. S. (2006). Use of facial composite systems in US law enforcement agencies. *Psychology, Crime & Law*, 12(5), 505-517. doi: 10.1080/10683160500254904
- Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: a meta-analytic review. *Psychology, Public Policy and Law*, 7(1), 3-35. doi: 10.1037/1076-8971.7.1.3
- Memon, A., Hope, L., & Bull, R. (2003). Exposure duration: effects on eyewitness accuracy and confidence. *British Journal of Psychology*, 94, 339-354. doi: 10.1348/000712603767876262
- Memon, A., Mastroberardino, S., & Fraser, J. (2008). Münsterberg's legacy: what does eyewitness research tell us about the reliability of eyewitness testimony?. *Applied Cognitive Psychology*, 22(6), 841-851. doi: 10.1002/acp.1487
- Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence-accuracy relationship for eyewitness identification decisions: effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied*, 19(1), 55-71.
- Pickel, K. L., Ross, S. J., & Truelove, R. S. (2006). Do weapons automatically capture attention? *Applied Cognitive Psychology*, 20(7), 871-893. doi: 10.1002/acp.1235
- Read, J. D. (1995). The availability heuristic in person identification - the sometimes misleading consequences of enhanced contextual information. *Applied Cognitive Psychology*, 9, 91-121. doi: 10.1002/acp.2350090202
- Reisberg, D., & Heuer, F. (2004). Memory for emotional events. In D. Reisberg & P. Hertel (Eds.), *Memory and emotion* (pp. 1- 41). New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195158564.003.0001
- Reisberg, D., & Heuer, F. (2007). The influence of emotion on memory in forensic settings. In M. P. Toglia, J. D. Read, D. F. Ross, & R. C. L. Lindsay (Eds.), *Handbook of eyewitness psychology* (Vol. 1, pp. 81-116). Mahwah, NJ: Laurence Erlbaum.
- Reynolds, J. K., & Pezdek, K. (1992). Face recognition memory: the effects of exposure duration and encoding instructions. *Applied Cognitive Psychology*, 6, 279-292. doi:10.1002/acp.2350060402
- Sauer, J. D., Brewer, N., Zuck, T., & Weber, N. (2010). The effect of retention interval on the confidence-accuracy relationship for eyewitness identification. *Law and Human Behavior*, 34, 337-347. doi:10.1007/s10979-009-9192-x
- Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin*, 100, 139-156. doi: 10.1037/0033-2909.100.2.139
- Stebay, N. M. (1992). A meta-analytic review of the weapon focus effect. *Law and Human Behavior*, 16(4), 413-424. doi: 10.1007/BF02352267
- Terry, R. L. (1993). How wearing eyeglasses affects facial recognition. *Current Psychology*, 12, 151-162. doi:10.1007/BF02686820
- Tollestrup, P., Turtle, J., & Yuille, J. (1994). Actual victims and witnesses to robbery and fraud: an archival analysis. In D. F. Ross, J. D. Read, & M. P. Toglia (Eds.), *Adult eyewitness testimony: current trends and developments* (pp. 144-160). New York, NY: Cambridge University Press.
- Wells, G. L. (1978). Applied eyewitness testimony research: system variables and estimator variables. *Journal of Personality and Social Psychology*, 36, 1546-1557. doi: 10.1037/0022-3514.36.12.1546
- Wells, G. L., & Hasel, L. E. (2007). Facial composite production by eyewitnesses. *Current Directions in Psychological Science*, 16, 6-10. doi: 10.1111/j.1467-8721.2007.00465.x
- Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology*, 54(1), 277-295. doi: 10.1146/annurev.psych.54.101601.145028

- Yuille, J. C., (1993). We must study forensic eyewitnesses to know about them. *American Psychologist*, 48, 572-573. doi:10.1037/0003-066X.48.5.572
- Yuille, J. C. (2013). The challenge for forensic memory research: Methodology. In B. S. Cooper, D. Griesel, & M. Ternes (Eds.), *Applied Issues in Investigative Interviewing, Eyewitness Memory, and Credibility Assessment* (pp. 3-18). Springer: New York.
- Zahradnikova, B., Duchovicova, S., & Schreiber, P. (2016). Facial composite systems. *Artificial Intelligence Review*, 1-22. doi:10.1007/s10462-016-9519-1
- Zaragoza, M. S., & Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20(4), 934-945. doi: 10.1037/0278-7393.20.4.934
- Renan Benigno Saraiva, Bacharel em Psicologia pela Universidade de Brasília (UnB), Brasil, é Doutorando em Psicologia na University of Portsmouth (UoP), Reino Unido. Endereço para correspondência: 69 Gains Road, Portsmouth, Hampshire, United Kingdom. PO40PJ. E-mail: renanbsaraiva@gmail.com
- Goiara Mendonça de Castilho, Doutora em Psicologia pela Universidade de Brasília (UnB), Brasil, é Professora Adjunto III na Universidade de Brasília (UnB). E-mail: goiaracastilho@gmail.com
- Raiane Nunes Nogueira, Bacharel em Psicologia pela Universidade de Brasília (UnB), Brasil. E-mail: rai.nogueira@gmail.com
- Letícia de Amorim Mota Coelho, Bacharel em Psicologia pela Universidade de Brasília (UnB), Brasil. E-mail: leticiaamorim.mota@gmail.com
- Luciana Carvalho Pulschen Alarcão, Bacharel em Psicologia pela Universidade de Brasília (UnB), Brasil. E-mail: lucianacpalarcao@gmail.com
- Jade Lage, Bacharel em Psicologia pela Universidade de Brasília (UnB), Brasil. E-mail: jadelagem@gmail.com

Recebido em 19.Jun.16
Revisado em 11.Jul.17
Aceito em 01.Ago.17